Cell, Vol. 109, 267–270, May 3, 2002, Copyright ©2002 by Cell Press

Specificity of Gene Regulation

Minireview

Beverly M. Emerson[1]
Salk Institute
Regulatory Biology Laboratory
10010 North Torrey Pines Road
La Jolla, California 92037

The physiologically coordinated expression of our genome requires exquisite regulation of gene specificity. Recent advances demonstrate that this formidable task is accomplished by diverse mechanisms and networks that operate at distinct levels within the nucleus.

The ability of each cell to program its genome and determine which genes are to be expressed at a given time and under specific stimuli is central to tissue differentiation, organogenesis, organismal development, and disease. The problem is that the same genetic information is contained within every cell, and many proteins that regulate gene activity can function nondiscriminately on DNA elements that control gene expression. Fortunately, a number of diverse mechanisms have evolved to ensure that the expression of our genome is a highly regulated process. A wealth of studies have shown that transcriptional activation occurs in discrete and controlled stages, from the packaging of a gene into chromatin and its localization within the nucleus to the recruitment of multiprotein complexes whose conformation and activity results from specific protein-protein or protein-DNA interactions. This minireview will focus on recent progress in elucidating the mechanisms by which regulated RNA polymerase II transcription is achieved at specific genes.

Targeted Chromatin Accessibility by Remodeling Complexes

A fundamental mechanism controlling the selectivity of gene expression is the limited ability of many transcription factors to access the genome. This is achieved by packaging genes into chromatin, which greatly impedes the binding of many proteins to their target DNA sequences. Accessibility of DNA to protein interaction is regulated by diverse enzymatic complexes that modulate nucleosomal structure by ATP-dependent "remodeling" or histone modification. Evidence for functional specificity of remodeling complexes came from early genetic analyses in yeast which demonstrated that mutations in different components of specific remodeling complexes, such as SWI/SNF, RSC, and ISWI 1, 2, produce distinct phenotypes. In addition, the *Drosophila* ISWI and BRM complexes were shown to localize in distinct regions of polytene chromosomes. Biochemical characterizations of numerous remodeling complexes have shown a diversity in protein composition that supported the idea of functional specificity. However, since all remodeling complexes can alter the structure of nucleosomes in general, little was known about how such complexes regulate genes specifically. Insight into this process was revealed in several recent studies. The demonstration that distinct chromatin-remodeling complexes function in a gene-specific manner was shown by analyzing transcriptional activation of nucleosomal templates by the *Drosophila* trxG protein Zeste, a DNA binding activator of homeotic genes (Kal et al., 2000). This work revealed that Zeste requires the BRM chromatin-remodeling complex for transactivation. The ISWI complex had no effect even though it efficiently modulates nucleosome structure. Thus, gene specificity is achieved by recruitment of BRM, but not ISWI, to Zeste bound promoters by direct interaction through the leucine zipper of Zeste and specific BRM subunits.

In another study, human SWI/SNF was shown to transcriptionally activate chromatin-assembled genes in a promoter-specific manner by selective association with particular classes of DNA binding proteins (Kadam et al., 2000). SWI/SNF was targeted to chromatinized β-globin promoters by zinc finger-containing proteins (ZFP) but not to HIV-1 promoters by either rel-containing NF-κB or the helix-loop-helix factor TFE-3. Although SWI/SNF transiently restructured nucleosomal DNA on both chromatin templates, stable remodeling only occurred when SWI/SNF was targeted by proteins with which it could physically interact. Interaction was mapped between the zinc finger DNA binding domains, rather than the activation domains, and the BRG1 ATPase subunit of SWI/SNF. In addition, a minimal complex composed of two recombinant subunits, BRG1 and BAF 155, was sufficient for targeted chromatin remodeling and transcriptional activation by ZFP in vitro. Mammalian SWI/SNF exists in a variety of biochemically diverse forms and a number of proteins, in addition to ZFP, have been shown to interact with these complexes through different subunits. One possibility is that particular domains of certain transcription factors can differentially target distinct SWI/SNF complexes to chromatin in a gene-selective manner. A specific form of SWI/SNF, called PBAF (polybromo- and BAF-containing complex), is distinguished from SWI/SNF by possessing a unique subunit, BAF180 (Xue et al., 2000). A recent study showed that PBAF, but not two other remodeling complexes (SWI/SNF and ACF), preferentially stimulates transcription by ligand-dependent nuclear hormone receptors, Sp1 and SREBP-1α (sterol response element binding protein), on chromatin-assembled target genes (Lemon et al., 2001).

Seminal experiments with distinct remodeling complexes from yeast and *Drosophila* also reveal specific targeting mechanisms by which selective gene regulation is achieved. For example, ySWI/SNF is recruited by acidic activation domains of transcription factors through direct interactions with the Snf5 subunit, which is not present in another related complex, yRSC (Boyer et al., 2000; Neely et al., 2002). The *Drosophila* ISWI-related complex, dNURF, associates with several different transcription factors through one of its four subunits, NURF301. Biochemical reconstitutions of dNURF reveal that NURF301 functions in a dual manner to catalyze nucleosome sliding with the ATPase ISWI and to direct

[1]Correspondence: emerson@salk.edu

promoter recruitment of NURF through specific factor interactions (Xiao et al., 2001). Interestingly, the yIswi2 complex is targeted to promoters by a transcriptional repressor, Ume6p, where it functions with a histone deacetylase complex to inactivate meiotic genes by generating a condensed, rather than accessible, nucleosomal structure (Goldmark et al., 2000).

Taken together, these studies suggest that chromatin-remodeling complexes function selectively through direct binding to specific transcription factor domains. In the absence of protein-directed recruitment, general nucleosome disruption catalyzed by remodeling enzymes may not be sufficient for stable gene activation. Similar mechanisms have been shown to exist for enzymatic complexes that modify histones: gene-targeted recruitment occurs by association of specific activators with particular histone acetyltransferases (HATs) or specific repressors with histone deacetylases (HDACs) (reviewed in Hassan et al., 2001a). The interplay between histone modification and recruitment of remodeling complexes is also a critical component of gene regulation. In fact, retention of SWI/SNF on nucleosomal promoters after recruitment by transcriptional activators is significantly stabilized by histone acetylation mediated by specific HAT complexes (Hassan et al., 2001b).

Nuclear Localization of Genes and Regulatory Proteins

The generation of chromatin structure that is capable of interacting with transcription factors is only one step in the process by which genes are specifically expressed. Another important step is nuclear organization, the position of genes within the nucleus and compartmentalization of proteins that regulate their expression. A recent study provided insight into this process by examining the activation of the β-globin gene locus during chemical induction of erythroid cell differentiation (Francastel et al., 2001). The erythroid-restricted heterodimeric protein NF-E2, composed of two subunits, p18 and p45, is a critical activator of differentiation and β-globin expression. Before differentiation NF-E2 is not bound to β-globin genes, and the inactive gene loci reside in the centromeric heterochromatin compartment. Interestingly, the two NF-E2 subunits are found in distinct nuclear compartments; p18 is in centromeric heterochromatin, whereas p45 is in the euchromatin compartment. During differentiation but prior to globin gene expression, p18 relocates to the euchromatin to join its partner, p45. Upon differentiation, the β-globin loci move from heterochromatin to euchromatin and are actively transcribed. Thus, β-globin gene activation is correlated with several linked processes involving concerted changes in nuclear organization. First, inactive genes are localized to condensed chromatin domains, and regulatory cofactors are sequestered in separate compartments until an appropriate signal for gene activation is received. Second, genes move to decondensed chromatin domains concomitant with protein relocalization to form functional complexes. Proteins and their target genes are apparently brought together in the same nuclear domain, where they associate and activate transcription.

Temporal Recruitment of Activators, Chromatin Enzymatic Complexes, and RNA Polymerase

The temporal order with which the multiple proteins that are required for transcriptional activation actually associate with their target promoters has been examined on several genes. One popular view is that activators must first bind to promoters, in conjunction with chromatin-remodeling or -modifying complexes, and then the core promoter recognition factors (TFIIA, B, D, E, F, and H plus RNAP II) are recruited to form an active initiation complex. However, recent results have indicated that the situation is much more complex, and the temporal order of factor-promoter binding can vary among different genes (reviewed by Fry and Peterson, 2001). For example, in the yeast HO promoter, the activator Swi5p initially binds and recruits the SWI/SNF remodeling complex followed by the SAGA HAT complex. These bound proteins then recruit a second activator, which engages the SRB mediator/coactivator followed by RNAP II. In the virally induced β interferon gene, the enhancer is nucleosome-free, but the TATA box is occupied by a positioned nucleosome. The HAT GCN5 is initially recruited and acetylates the nucleosome, which, in turn, enables a complex containing the HAT, CBP, and RNAP II holoenzyme to associate with the initiation region. A critical step in this process is the recruitment of SWI/SNF by CBP, which remodels the phased nucleosome and facilitates TFIID binding and transcription.

A recent study examined the order in which activators, initiation factors, and chromatin enzymatic complexes assembled on the human α1 antitrypsin promoter during cellular differentiation (Soutoglou and Talianidis, 2002). Surprisingly, they found that the promoter was initially bound by an activator, HNF-1α, and two components of the initiation complex, TFIIB and TATA binding protein (TBP) within a positioned nucleosome, at the beginning of the differentiation process long before transcriptional activation. As differentiation progressed, other initiation factors, including RNAP II, were recruited, and a complete TFIID complex replaced TBP. Subsequently, the HAT complexes CBP and PCAF associated, which resulted in histone hyperacetylation within the proximal promoter. At the time of transcription, a second activator, HNF-4, bound and SWI/SNF interacted transiently to remodel the nucleosome encompassing the TATA box, which coincided with the release of RNAP II. Thus, the assembly of protein complexes that regulate the α1 antitrypsin promoter occurs in an intricately staged manner during differentiation, and initiation factors are present at very early times.

It is important to consider the chromosomal context in which a particular gene resides when analyzing the requirements and timing of protein binding and chromatin-remodeling or -modifying events. For example, the α1 antitrypsin promoter resides within a 100 kb domain containing a cluster of genes and numerous expression-related DNase hypersensitive sites (DHSs), which often reflect chromatin-remodeling events through protein-nucleosome interactions. Expression of the activators HNF-1 and HNF-4 induce multiple DHSs within this cluster (Rollini et al., 1999). Therefore, the order of events that correlate with α1 antitrypsin transcription may differ from that of other DHSs within the locus and may be strongly influenced by the dynamic changes in chromatin structure and protein recruitment occurring at neighboring genes. The ability of "poised" initiation complexes to form at a positioned nucleosome within the α1 antitrypsin but not the β interferon promoter is unclear but may reflect the occurrence of low-level, basal

transcription prior to full activation of the α1 antitrypsin gene.

Tissue-Selective Initiation Factors and Induced Conformations of Coactivator Proteins

The idea that transcriptional specificity is conferred by tissue-restricted activators that relay their effects through common core promoter recognition factors has been challenged recently. In one interesting analysis, the subunit composition of the initiation factor TFIID, consisting of TBP and TAFII components, had a profound effect on organogenesis (Freiman et al., 2001). Specifically, the loss of one particular subunit, TAFII105, resulted in defective ovarian development in mice. This was correlated with misexpression of a subset of genes and inappropriate inhibin-activin signaling. Thus, tissue-selective TAF subunits can exist within diverse TFIID complexes as part of the "general" transcription machinery and can contribute significantly to proper patterns of gene expression, which are necessary for specialized cell function.

Transcription not only requires binding to the promoter of activator proteins, RNA polymerase, and initiation factors but also to proteins called "mediators." These large complexes interact with either activators or repressors and RNAP II to facilitate or repress transcription by communicating signals from regulatory factors to the basal initiation machinery. In an elegant biochemical analysis, the structures of two mammalian mediator complexes, ARC-L and CRSP, were determined using electron microscopy (Taatjes et al., 2002). One complex, ARC-L, is transcriptionally inactive whereas CRSP is highly active. The two complexes are structurally dissimilar, yet each one displays a high degree of conformational flexibility. Interestingly, CRSP undergoes pronounced structural changes when bound to different transcriptional activators. Interaction of CRSP with VP16 or SREBP through separate regions of the coactivator generates distinct, extended protein conformations. This suggests that CRSP function may be determined by the exact structural conformation it assumes when bound by distinct activators. Such structural variability may enable CRSP to regulate gene expression selectively by relaying activation signals to distinct components of the RNAP II machinery within specific promoter contexts.

Switching of Transcriptional Programs

Cell differentiation and response to signaling pathways require that certain subsets of genes be activated while other genes are repressed. One can envision many scenarios in which this is achieved. Importantly, recent studies have defined novel and distinct mechanisms through which such transcriptional switches occur. One study analyzed the determinants of the transcriptional switch between two distinct signaling pathways controlled by nuclear hormones or cAMP (Xu et al., 2001). Surprisingly, this switch is mediated by the dual ability of the CARM1 methyltransferase to methylate the cofactor CBP/p300 or nucleosomes when targeted by nuclear hormone receptors. Methylation of CBP/p300 prevents it from being recruited to promoters bound by the activator CREB, which then fails to stimulate cAMP-dependent transcription. CARM1 binds directly to CBP/p300 or nuclear receptors, but not to CREB. In this way, CARM1

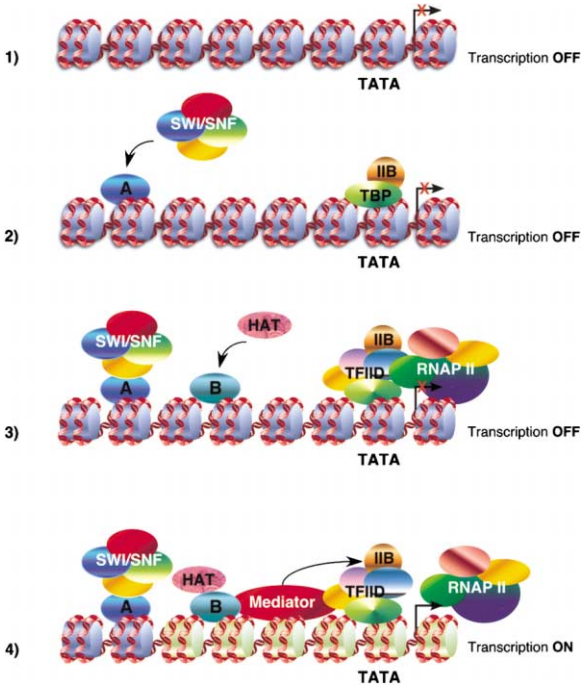

Figure 1. Model of Transcriptional Activation

This diagram outlines potential steps in the generation of a transcriptionally active promoter in chromatin. (1) The promoter is assembled into a nucleosomal structure that is transcriptionally inactive. (2) Interaction of a sequence-specific DNA binding protein (A) recruits a chromatin-remodeling complex (SWI/SNF), which results in stabilized binding of protein A through an ATP-dependent perturbation of nucleosomal structure. In some promoters, a partial initiation complex (TFIIB, TBP) may also be bound at this stage. (3) After remodeling, a histone acetyltransferase complex (HAT) is targeted to the promoter, where it acetylates nucleosomes (in green) and facilitates the binding of a second transcriptional activator (B). A complete initiation complex (TFIID, RNAP II) may be formed at this stage on some promoters. (4) Protein B engages a mediator/coactivator complex (Mediator) and induces a particular structural conformation, which imparts specificity to its interaction with components of the initiation complex. This results in RNAP II release and activated transcription.

The requirement for chromatin-remodeling events and specific histone modifications varies among different promoters as well as the exact timing of initiation complex formation. The adult β-globin promoter apparently undergoes remodeling and interacts with most of its regulatory proteins in the heterochromatic compartment before relocating to euchromatin during erythroid differentiation. This relocalization is correlated with association of a particular activator (NF-E2) and gene expression.

functions as a coactivator of hormone-inducible promoters and a corepressor of cAMP-responsive genes through nucleosome or cofactor methylation.

Another study focused on the pituitary-specific POU domain factor Pit-1, which activates growth hormone gene expression in somatotropes and represses its expression in lactotropes (Scully et al., 2000). The Pit-1 DNA binding sites were found to function allosterically by virtue of a 2-base pair spacing that facilitates the bipartite structure of the interacting POU domain. Conformational changes of Pit-1 bound to this site promote the recruitment of corepressor complexes, such as N-CoR, that inactivate the growth hormone gene in lactotropes. Thus, DNA allosterism modulates the confor-

mation of a bound protein, which, in turn, influences the nature of the interacting cofactors. This mode of regulation is likely to be critical for many transcription factors that interact with structurally diverse DNA sequences, such as the tumor suppressor protein p53. Finally, a recent study revealed that naturally occurring isoforms of the LEF-1/TCF enhancer factor, which have distinct roles in activation or repression of Wnt signaling in vivo, display dramatically different intrinsic affinities for chromatin despite the presence of identical DNA binding domains (Tutter et al. 2001). In particular, full-length LEF-1 protein was shown to bind weakly to chromatin but cooperatively with its transcriptional coactivator β-catenin. By contrast, dominant-negative forms of LEF-1 lacking the β-catenin interaction domain bound chromatin avidly and independently of the coactivator. Thus, cooperative interactions with coactivator and corepressor complexes on chromatin can also regulate enhancer complex assembly and function.

Perspectives

Together these studies illustrate some of the elegant mechanisms by which selective gene expression is achieved within our complex genome (for a model of transcriptional activation, see Figure 1). Rapid advances in technology will continue to reveal exciting new mechanisms as they emerge. Key unresolved questions include defining the basis for gene-specific targeting of chromatin enzymatic activities and understanding how these complexes are deregulated or mistargeted in human disease (Di Croce et al., 2002). Detecting macromolecular interactions in living cells in real time and performing crystallographic analyses of large multisubunit complexes will reveal the changes in protein compositions and structural conformations that we have seen are critical determinants of gene specificity. Mapping global patterns of protein binding and epigenetic modifications of chromatin and DNA will also be invaluable to elucidate how genomic programming is normally established and how it changes with disease. Certainly a challenging new frontier will be to decipher the principles of nuclear organization and to analyze how gene expression is controlled through dynamic changes in chromosome and protein localization. It will be especially enlightening to learn whether temporal protein recruitment to genes actually occurs in distinct nuclear compartments. Although it is clearly a daunting task to understand the regulated expression of our entire genome, detailed mechanistic studies at the level of individual genes affords us the capacity to develop the paradigms necessary for a truly comprehensive view of genome regulation within the framework of our complex physiology.

Selected Reading

Boyer, L.A., Logie, C., Bonte, E., Becker, P.B., Wade, P.A., Wolffe, A.P., Wu, C., Imbalzano, A.N., and Peterson, C.L. (2000). J. Biol. Chem. *275*, 18864–18870.

Di Croce, L., Raker, V.A., Corsaro, M., Fazi, F., Fanelli, M., Faretta, M., Fuks, F., Lo Coco, F., Kouzarides, T., Nervi, C., et al. (2002). Science *295*, 1079–1082.

Francastel, C., Magis, W., and Groudine, M. (2001). Proc. Natl. Acad. Sci. USA *98*, 12120–12125.

Freiman, R.N., Albright, S.R., Zheng, S., Sha, W.C., Hammer, R.E., and Tjian, R. (2001). Science *293*, 2084–2087.

Fry, C.J., and Peterson, C.L. (2001). Curr. Biol. *11*, R185–R197.

Goldmark, J.P., Fazziom, T.G., Estep, P.W., Church, G.M., and Tsukiyama, T. (2000). Cell *103*, 423–433.

Hassan, A.H., Neely, K.E., Vignali, M., Reese, J.C., and Workman, J.L. (2001a). Front. Biosci. *6*, D1054–D1064.

Hassan, A.H., Neely, K.E., and Workman, J.L. (2001b). Cell *104*, 817–827.

Kadam, S., McAlpine, G.S., Phelan, M.L., Kingston, R.E., Jones, K.A., and Emerson, B.M. (2000). Genes Dev. *14*, 2441–2451.

Kal, A.J., Mahmoudi, T., Zak, N.B., and Verrijzer, C.P. (2000). Genes Dev. *14*, 1058–1071.

Lemon, B., Inouye, C., King, D.S., and Tjian, R. (2001). Nature *414*, 924–928.

Neely, K.E., Hassan, A.H., Brown, C.E., Howe, L., and Workman, J.L. (2002). Mol. Cell. Biol. *22*, 1615–1625.

Rollini, P., Xu, L., and Fournier, R.E. (1999). Somat. Cell Mol. Genet. *25*, 207–221.

Scully, K.M., Jacobson, E.M., Jepsen, K., Lunyak, V., Viadiu, H., Carriere, C., Rose, D.W., Hooshmand, F., Aggarwal, A.K., and Rosenfeld, M.G. (2000). Science *290*, 1127–1131.

Soutoglou, E., and Talianidis, I. (2002). Science *295*, 1901–1904.

Taatjes, D.J., Naar, A.M., Andel, F., 3rd, Nogales, E., and Tjian, R. (2002). Science *295*, 1058–1062.

Tutter, A.V., Fryer, C.J., and Jones, K.A. (2001). Genes Dev. *15*, 3342–3354.

Xiao, H., Sandaltzopoulos, R., Wang, H.-M., Hamiche, A., Ranallo, R., Lee, K.-M., Fu, D., and Wu, C. (2001). Mol. Cell *8*, 531–543.

Xu, W., Chen, H., Du, K., Asahara, H., Tini, M., Emerson, B.M., Montminy, M., and Evans, R.M. (2001). Science *294*, 2507–2511.

Xue, Y., Canman, J.C., Lee, C.S., Nie, Z., Yang, D., Moreno, G.T., Young, M.K., Salmon, E.D., and Wang, W. (2000). Proc. Natl. Acad. Sci. USA *97*, 13015–13020.